TOWERS PERRIN

TILLINGHAST

Casualty Loss Reserve Seminar

The Language of Uncertainty:

Terminology Surrounding Loss Reserve Variability/Ranges

Dan Murphy

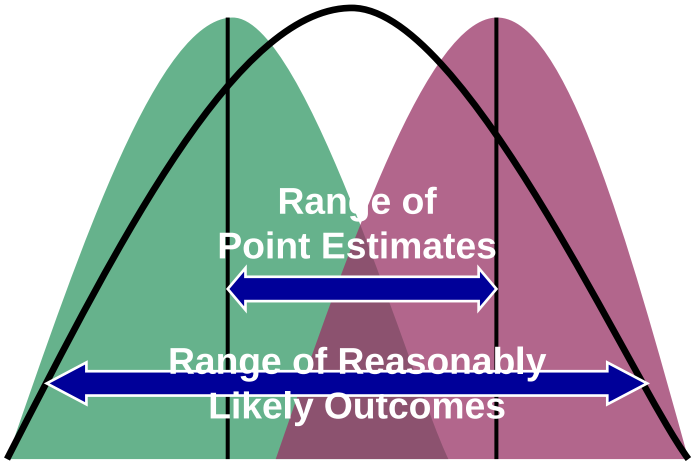
September 2007

Why analyze the variability of claim liabilities?

NAIC	■ Actuarial opinions are produced on a "reasonableness" standard ■ Variation from the "best estimate" is the issue ■ Actuarial Opinion Summary (AOS) includes focus on ranges
SEC	■ Require public companies to discuss reserve uncertainty in 10-K filings ■ Increasing pressure…hand-waving rationale will soon be inadequate
Rating Agencies	■ Capital adequacy analyses usually assume reserve shortfalls ■ Management is expected to consider more than just the best estimate
Actuarial Profession	■ CAS Task Force recommends greater focus on advanced techniques for measuring variability of reserves ■ UK group reaches similar findings
Fiduciary Duty	■ It is prudent business practice to recognize your business risks to the best of your ability ■ Stochastic analysis can provide strategic operational and financial insights

What is an estimate of claim liabilities?

- A point estimate of the ultimate value of outstanding claim liabilities is a prediction of the mean value of that random variable
- We do not attempt to predict the amount by which ultimate liabilities will deviate from the “true mean”
- Where can our estimate of the ultimate value of outstanding claim liabilities go wrong relative to the eventual value of outstanding claims?



  - Our estimate of the mean may not coincide with the true mean
  - Ultimate liabilities will ultimately deviate from their true mean

A “real life” analogy

- Rather than estimating the value of future payments going out of the claims department, suppose we want to estimate the IQ of the next person to walk in the back door
- If we knew the average IQ of all CLRS registrants, that would be our guess
- Alternatively, our method is to measure the average IQ of everyone in Room ___X___
  - … everyone take out a piece of paper …
- When Joe walks in, by what amount could our guess be wrong?
  - The average IQ in the room could be different from true CLRS mean IQ
  - Joe’s IQ could be different from the true mean

Several distinct types of risks are inherent in the estimation of claim liabilities

Total Risk

Estimation error
Difference between estimated mean and true mean

Parameter Risk
Variability due to fact that our model's parameters are estimates

Model Risk
Variability due to fact that our model imperfectly represents reality

Process Risk
Difference between actual costs and true mean

Our room's average IQ today could be different from the average of all such samples from Room __X__

Room X's population may not be representative of the entire CLRS population: IQ biased on the high side?

The IQ of the Joe walking through the door will differ from the true CLRS mean

Estimate of Expected Outcome

Increase or decrease with volume?

True Mean Outcome

Increase or decrease with volume?

Eventual Outcome

Total Risk, aka Mean Square Error, is the statistical equivalent of the Pythagorean Theorem

Mean of Potential Deterministic Estimates = Liability's True Mean (if unbiased)

Method's Estimate of Accident Year i's Outstanding Liability

$\hat{C}_i$

Parameter Risk

"Explained Deviation"

bias

$E(\hat{C}_i) = E(C_i)$

(if unbiased)

Process Risk

"Unexplained Deviation"

Total Risk

aka, Mean Square Error

$$mse(\hat{C}_i) = E(\hat{C}_i - C_i)^2$$

$$= Var(\hat{C}_i) + Var(C_i)$$

excluding bias

C_i

Accident Year i's Eventual Outstanding Liability

One last page of definitions

- Sometimes the term “risk” refers to variance and sometimes to standard deviation, depending on the context
- Most of the time “mean square error” and “total variance” are used interchangeably
  - When the estimate is unbiased, mse and total variance are identical
  - What is bias?
    - — Bias occurs when the mean value of your estimate/estimation procedure of ultimate loss does not coincide with the true mean value of ultimate loss for some “systematic” reason not captured by your procedure
- Estimation error = parameter risk + bias^2
- Coefficient of Variation, or CV
  - Coefficient of Variation (X) = Standard Deviation (X) ÷ Mean (X)
  - Note: Standard Deviation (Ultimate Loss) =
Standard Deviation (Outstanding Loss) =
Standard Deviation (IBNR)
because the only difference between the various quantities is a known dollar amount
  - Stochastic methods generally estimate the standard deviation by analyzing estimates of ultimate loss
  - The CVs of ultimate loss, unpaid loss, and IBNR will differ because, although their numerators are the same, their denominators are different

Relevant sources of variability depend on the question at hand

- When is it important to analyze total risk?
  - Financial solvency/economic capital context
  - When solvency is the issue ➡ all sources of risk are relevant
- When is it important to analyze only parameter risk?
  - "Reasonable range of estimates"
  - When the reasonability of the estimate is the issue ➡ only estimation error (= parameter risk in the absence of bias) is relevant
- Ability to estimate those separate sources of risk varies by stochastic method
  - Mack ➡ both parameter and process risk estimation are built-in
  - Bootstrap ➡ measures parameter risk; process risk measurement can be an "overlay"
  - Monte Carlo methods ➡ e.g., the "process" of loss development factors is simulated around a selected factor; the variability of the factor – the parameter – must be handled separately

Mack Method

Bootstrapping

Practical Method

Mack Method: Overview

- Mack Method derives formulas for the standard error of the reserves projected by the chain ladder method
- Tillinghast recommends using the recursive formulas from Murphy's 1994 paper "Unbiased Loss Development Factors"
- The formulas provide for process and parameter risk, separately and in total
- The method can be extended to incorporate age-to-age factors other than the volume weighted average
- A normal or lognormal distribution is fit to the mean and variance of the reserve to yield a distribution of reserves
- The variability of the tail beyond the triangle can be incorporated in various ways

Mack: Summary

Advantages	Disadvantages

■ Widely regarded in the industry
■ Founded in statistical theory
■ Works with chain-ladder eligible triangles
■ Can reflect tail variability
	■ Data outliers can have a leveraged effect on the results
	■ May over-parameterize the risk
	  ■ n(n+1)/2 data points
	  ■ 2(n-1) parameters: link ratios and the development periods' variances

Mack Method

Bootstrapping

Practical Method

Bootstrap Method: Overview

- Bootstrapping is a simulation technique that generates empirical probability distributions of complex functions
- A triangle of cumulative fitted values for the past triangle is obtained by backwards recursion on the most recent diagonal using standard chain ladder link ratios
- A set of Pearson residuals is calculated from the fitted and actual data
- Bootstrapping utilizes the sampling-with-replacement technique on the residuals of the historical data
- Each simulated sampling scenario produces a new “realization” of triangular data that has the same statistical characteristics as the actual data
- Our model calculates both parameter and total risk
- Our bootstrapping implementation can calculate tail volatility by employing curve fitting to each realization of average loss development factors
- Our Bootstrapping implementation includes a B-F option for the new data “realization”
- Outlier observations can be restricted
- The sampling of residuals can be restricted for the first development period

Bootstrapping: Summary

Advantages	Disadvantages
■ Easy to understand and explain ■ Commonly used in industry ■ Accommodates BF method ■ Facilitates the calculation of tail volatility	■ Data outliers can have a leveraged effect on the results ■ Method does not work well with negative loss development (due to underlying theoretical model)

Mack Method

Bootstrapping

Practical Method

Practical Method: Basic Theory and assumptions

- The Practical Method uses Monte Carlo simulation to estimate reserve distributions based on the three most popular deterministic methods – Chain Ladder, Loss Ratio, and Bornhuetter-Ferguson
- Practical simulates age-to-age (ATA) factors as normal or lognormal r.v.'s and loss ratios as normal r.v.'s.
  - Means and variances of those distributions are selected inputs
  - For BF method, LDFs can be "fixed" based on the ATA means, or "variable" based on the ATA simulations
  - The variability of the tail factor is a manual entry or, if left blank, can be modelled by assuming that the standard error of the last age-to-age factor is repeated for as many years as the user selects
- Explicitly reflects process risk only, but parameter can be incorporated judgmentally

Practical: Summary

Advantages	Disadvantages
■ Easy to understand and explain ■ Accommodates the three most popular actuarial deterministic methods ■ Can incorporate tail variability	■ Not as well known in the actuarial community ■ Does not explicitly measure parameter risk

Flexibility allows user to obtain wide range of results.

Tail variability

- Many of the popular stochastic methods only measure risk to the edge of the triangle
  - Variability for development beyond the triangle – so called “tail variability” – must be measured and incorporated separately
- Mack suggested a heuristic approach to tail variability in his 1999 paper
- Bootstrap (per England and Verrall, 1998) also only measures risk to the edge of the triangle
- Monte carlo simulation can simulate a tail assuming you have an idea of the variability of your tail factor
- Imputed Variability is an alternative tail variability approach that could be used to supplement any stochastic method whose variability estimate is to the edge of the triangle

Method I: Tail variability can be reflected with the Mack Method using the heuristic in his 1999 paper ...

Mack Method **Taylor-Ashe Data**

Taylor, G., and F. Ashe, "Second Moments of Estimates of Outstanding Claims," *Journal of Econometrics*, **1983**, 23, pp. 37-61.

AY/DY	1	2	3	4	5	6	7	8	9	10	Ultimate
i/k	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=∞
i=1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463	4,291,609
i=2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	5,433,719	5,977,091
i=3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315	5,285,148	5,378,826	5,916,709
i=4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268	4,835,458	5,205,637	5,297,906	5,827,696
i=5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311	4,207,459	4,434,133	4,773,589	4,858,200	5,344,020
i=6	396,132	1,333,217	2,180,715	2,985,752	3,691,712	4,074,999	4,426,546	4,665,023	5,022,155	5,111,171	5,622,289
i=7	440,832	1,288,463	2,419,861	3,483,130	4,088,678	4,513,179	4,902,528	5,166,649	5,562,182	5,660,771	6,226,848
i=8	359,480	1,421,128	2,864,498	4,174,756	4,900,545	5,409,337	5,875,997	6,192,562	6,666,635	6,784,799	7,463,279
i=9	376,686	1,363,294	2,382,128	3,471,744	4,075,313	4,498,426	4,886,502	5,149,760	5,544,000	5,642,266	6,206,493
i=10	344,014	1,200,818	2,098,228	3,057,984	3,589,620	3,962,307	4,304,132	4,536,015	4,883,270	4,969,825	5,466,807
sum below diagonal	0	1,200,818	4,480,356	10,704,484	16,654,155	22,458,247	28,603,165	34,979,600	42,942,618	49,137,483	58,342,840
										selected	
f_i = ATAs	3.491	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018	1.100	
CDFs	15.891	4.553	2.605	1.788	1.523	1.380	1.270	1.205	1.119	1.100	
										heuristic	
σ_i^2	160,280	37,737	41,965	15,183	13,731	8,186	447	1,147	447	13,731	
σ_f^2	0.04817	0.00368	0.00279	0.00082	0.00076	0.00051	0.00004	0.00013	0.00012	0.00076	

… Variance estimates are completed for future development periods using the Murphy formulas …

Mack Method **Taylor-Ashe Data**

Total Variance of Chain Ladder Projection

i/k	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=∞	Standard Error
i=1											6.52E+10	255,358
i=2										5.71E+09	1.04E+11	322,627
i=3									8.88E+09	1.48E+10	1.14E+11	337,491
i=4								2.79E+09	1.19E+10	1.78E+10	1.16E+11	340,279
i=5							3.94E+10	4.63E+10	6.14E+10	6.83E+10	1.67E+11	409,208
i=6						6.11E+10	1.14E+11	1.29E+11	1.58E+11	1.69E+11	2.95E+11	542,772
i=7					6.29E+10	1.46E+11	2.19E+11	2.46E+11	2.95E+11	3.12E+11	4.80E+11	692,462
i=8				1.43E+11	2.75E+11	4.21E+11	5.56E+11	6.21E+11	7.32E+11	7.66E+11	1.06E+12	1,027,513
i=9			5.83E+10	2.40E+11	3.93E+11	5.47E+11	6.93E+11	7.73E+11	9.05E+11	9.44E+11	1.24E+12	1,115,196
i=10		6.08E+10	2.36E+11	6.03E+11	8.84E+11	1.14E+12	1.38E+12	1.54E+12	1.79E+12	1.86E+12	2.34E+12	1,528,545
Sum		6.08E+10	3.07E+11	1.11E+12	1.91E+12	2.92E+12	4.02E+12	4.52E+12	5.50E+12	5.99E+12	1.03E+13	3,202,424
Standard Error		246,656	553,840	1,053,717	1,382,465	1,710,220	2,005,746	2,126,418	2,345,361	2,447,618	3,202,424	

… The 10% tail increases total CV of the reserve by 20 only basis points

Mack Method **Taylor-Ashe Data**

AY	Est'd Ultimate	Liability	Standard Errors w/ Murphy Recursive Formulas: Process risk	Parameter risk	Total risk	CV: Process risk	Parameter risk	Total risk
1	4,291,609	390,146	231,457	107,868	255,358	59.3%	27.6%	65.5%
2	5,977,091	638,006	278,384	163,066	322,627	43.6%	25.6%	50.6%
3	5,916,709	1,007,394	289,437	173,570	337,491	28.7%	17.2%	33.5%
4	5,827,696	1,239,428	292,387	174,067	340,279	23.6%	14.0%	27.5%
5	5,344,020	1,470,709	359,922	194,697	409,208	24.5%	13.2%	27.8%
6	5,622,289	1,930,577	482,479	248,627	542,772	25.0%	12.9%	28.1%
7	6,226,848	2,743,718	616,826	314,688	692,462	22.5%	11.5%	25.2%
8	7,463,279	4,598,781	916,626	464,305	1,027,513	19.9%	10.1%	22.3%
9	6,206,493	4,843,199	1,023,695	442,392	1,115,196	21.1%	9.1%	23.0%
10	5,466,807	5,122,793	1,437,309	520,187	1,528,545	28.1%	10.2%	29.8%
Total:	58,342,840	23,594,604	2,235,431	2,293,113	3,202,424	9.5%	9.7%	13.6%

With no tail

Total:	53,038,946	18,290,709	1,878,292	1,569,349	2,447,618	10.3%	8.6%	13.4%

Diff.:	5,303,895	5,303,895	357,139	723,764	754,806	-0.8%	1.1%	0.2%

Other stochastic methods can be expanded to incorporate tail variability: two examples

- Bootstrap
  - For each simulated "false history" triangle, fit a power curve to the link ratios that result
  - The ultimates in each simulation will reflect different tails
- Practical (Tillinghast's version of monte carlo stochastic reserving)
  - Assume you have a tail from another source whose distribution you want to simulate in a monte carlo fashion
  - Your tail is the mean of the distribution
  - For the distribution's variance, you can try
    - the Mack heuristic or
    - the Imputed Tail Variability technique below

Imputed Tail Variability: Using selected ultimates by accident year, calculate the implied tail factor variability

XYZ Insurance: Workers Compensation

Accident Period	Selected Ultimate Loss & ALAE	Reported Loss Dev to 120 Mos.	Implied Rptd Tail Factors
1997	254,354,038	251,134,948	1.01282
1998	242,595,872	239,481,279	1.01301
1999	247,896,393	244,969,955	1.01195
2000	261,515,841	258,477,171	1.01176
2001	293,849,905	290,199,024	1.01258
2002	288,083,714	283,993,986	1.01440
2003	325,907,056	321,275,467	1.01442
2004	382,581,046	375,166,184	1.01976
2005	431,081,736	409,163,650	1.05357
2006	450,124,774	400,874,277	1.12286
Tot/Wtd Avg	3,177,990,375	3,074,735,942	1.03358
Process Variance ≈ var(AY link ratios)			0.001282
Parameter Variance (Process variance / n-1)			0.000142

- Each accident year gives rise to a different estimated 120-ult tail factor = f_{10}
- The formula for the parameter variance is the well-known formula for the variance of a population mean
- Weighted averages were used here; if the simple average is selected, use Excel's 'var' function
- Use formulas for the variance of the product of independent random variables (next page)

The imputed variance of the 1.034 tail adds 260 basis points to the CV of estimated loss

Process risk: $Var(C_{10} \cdot f_{10}) = Var(C_{10})E^2(f_{10}) + Var(f_{10})E^2(C_{10}) + Var(C_{10})Var(f_{10})$

Parameter risk: $Var(\hat{C}_{10} \cdot \hat{f}_{10}) = Var(\hat{C}_{10})E^2(\hat{f}_{10}) + Var(\hat{f}_{10})E^2(\hat{C}_{10}) + Var(\hat{C}_{10})Var(\hat{f}_{10})$

XYZ Insurance: Workers Compensation	Risk		
	Parameter	Process	Total
Estimated loss developed to 120 mos (prior slide)	3,074,735,942	3,074,735,942	3,074,735,942
Variance of total loss developed to 120 mos	2.4219E+14	1.0672E+15	1.3094E+15
(Mack calculation not shown)			
CV of developed loss to 120 mos	0.5%	1.1%	1.2%
Mean Tail Factor (prior slide)	1.034	1.034	1.034
Variance of tail factor (prior slide)	0.0001424	0.0012818	
Var(developed loss to 120 mos x tail) (formulas above)	1.6052E+15	1.3259E+16	1.4864E+16
Standard error	40,064,600	115,148,335	121,919,282
Estimated IBNR **(not shown)**	380,749,837	380,749,837	380,749,837
CV of estimated IBNR	10.5%	30.2%	32.0%
Estimated Ultimate Loss (prior slide)	3,177,990,375	3,177,990,375	3,177,990,375
CV of Estimated Ultimate Loss	1.3%	3.6%	3.8%

Aggregation techniques: across lines

- Means aggregate without much fuss:
  - E(X+Y) = E(X) + E(Y)

    i.e., having the marginal moment gives you the aggregate moment
- Variances aggregate without much fuss when the lines are independent – er, too strong – uncorrelated
  - Var(X+Y) = Var(X) + Var(Y)

    i.e., having the marginal variance gives you the aggregate variance iff X and Y are not correlated
- If you suspect that your lines' are skewed, and the first two moments won't suffice to capture the full shape, you will have to resort to "other means"
  - The solution du-jour utilizes "Copulas"

Refresher points on covariance and correlation

- Covariance is to the formula for the variance of the sum of two random variables as the cross product term is to the square of a binomial

$$Var(X \pm Y) = Var(X) \pm 2Cov(X,Y) + Var(Y)$$

- Correlation scales the two lines by dividing by their standard deviations

$$\text{Correlation Coefficient}: \quad \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

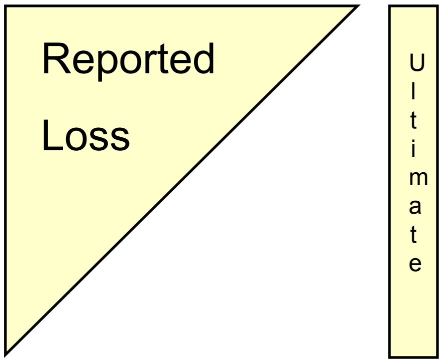
  - Correlation is “standardized” covariance

- The concepts are expanded to a vector of distributions X for N lines of business via the covariance matrix and the correlation matrix

$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & \cdots & Cov(X_1,X_N) \\ Cov(X_2,X_1) & Var(X_2) & \cdots & Cov(X_2,X_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_N,X_1) & Cov(X_N,X_2) & \cdots & Var(X_N) \end{bmatrix} \quad Corr = \begin{bmatrix} 1 & corr(X_1,X_2) & \cdots & corr(X_1,X_N) \\ corr(X_2,X_1) & 1 & \cdots & corr(X_2,X_N) \\ \vdots & \vdots & \ddots & \vdots \\ corr(X_N,X_1) & corr(X_N,X_2) & \cdots & 1 \end{bmatrix}$$

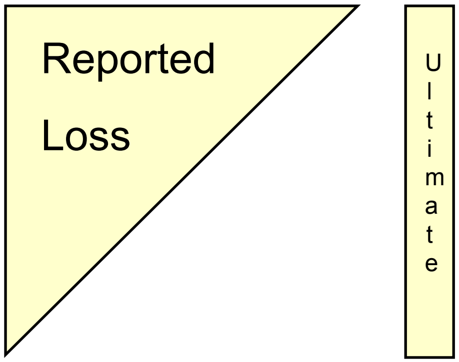
  - These matrices are positive-semidefinite (i.e., you can “square root” them, analogous to a standard deviation and a variance)
  - Can be inverted only if positive-definite (cannot be “zero”)

Correlation between lines measures the degree to which deviations from the mean move – or don’t move – in conjunction with each other



Auto BI

$$C_{ABI}(i,J)-\bar{C}_{ABI}(i,J)$$



GL

$$C_{GL}(i,J)-\bar{C}_{GL}(i,J)$$

- Correlation is a linear concept
- Given pairwise samples of ultimates from two lines for I accident years, the strength to which the lines “co-vary” is estimated by the statistic

$$\frac{1}{I-1}\sum_{i=1}^{I}\left(C_{ABI}(i,J)-\bar{C}_{ABI}(i,J)\right)\left(C_{GL}(i,J)-\bar{C}_{GL}(i,J)\right)$$

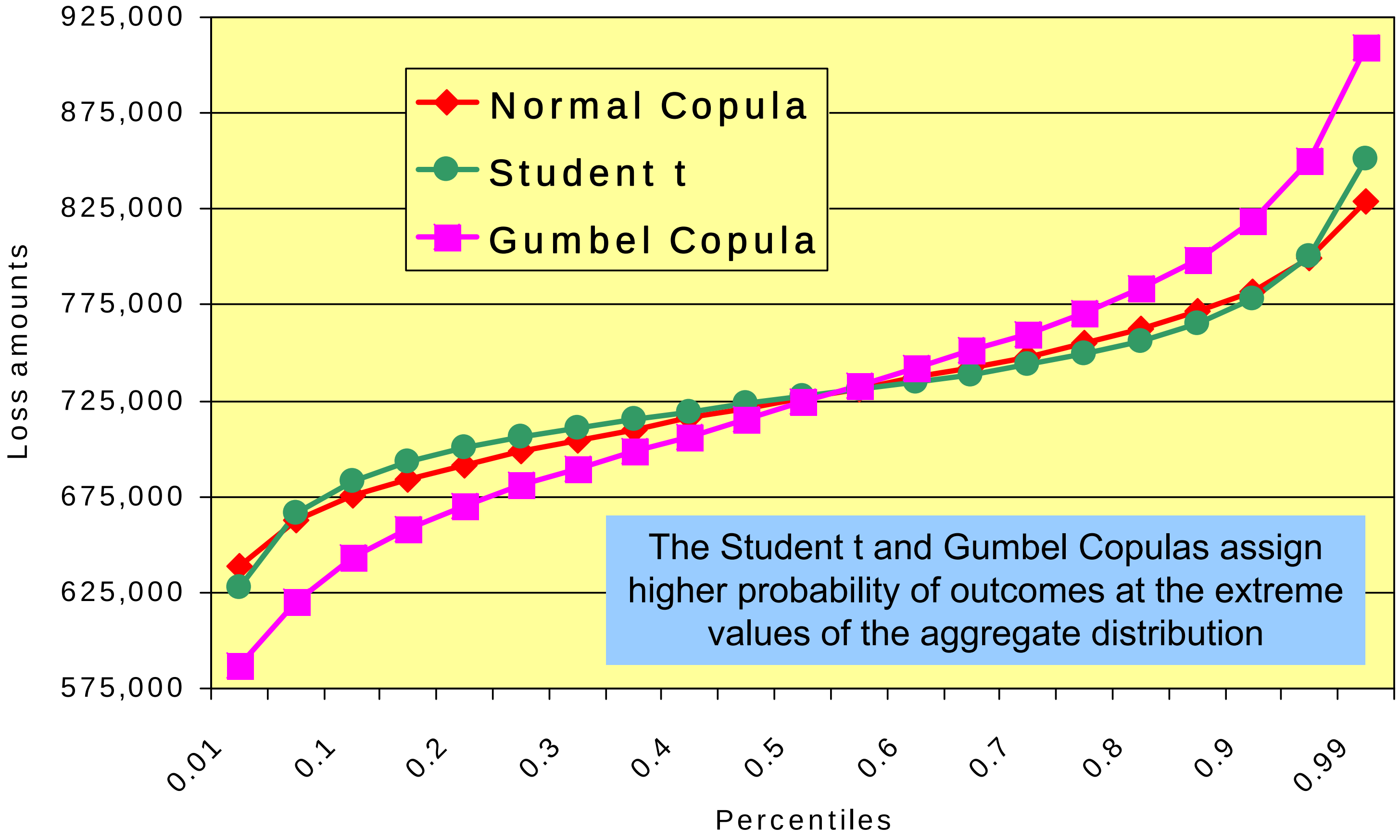
Correlation measures the average strength of association between two lines

- Correlation can help you find another moment of the aggregate distribution of two or more lines if you intend to model aggregate losses with a two parameter probability distribution (e.g., lognormal)
- But a two parameter probability distribution is insufficient to model a joint distribution whose strength of association increases in the tail
  - Example: for property losses, the correlation is higher in the tail of the distribution
- Ideally one would like to specify the complete joint distribution of N lines of business
- It turns out that every joint distribution of N lines of business can be decomposed into its N marginal distributions by virtue of a amalgamating function called a "copula"
- Vice versa, given the marginal distributions of N lines of business, the joint distribution can be calculated with the appropriate copula

Copulas provide a convenient way to express the aggregate distribution of several lines

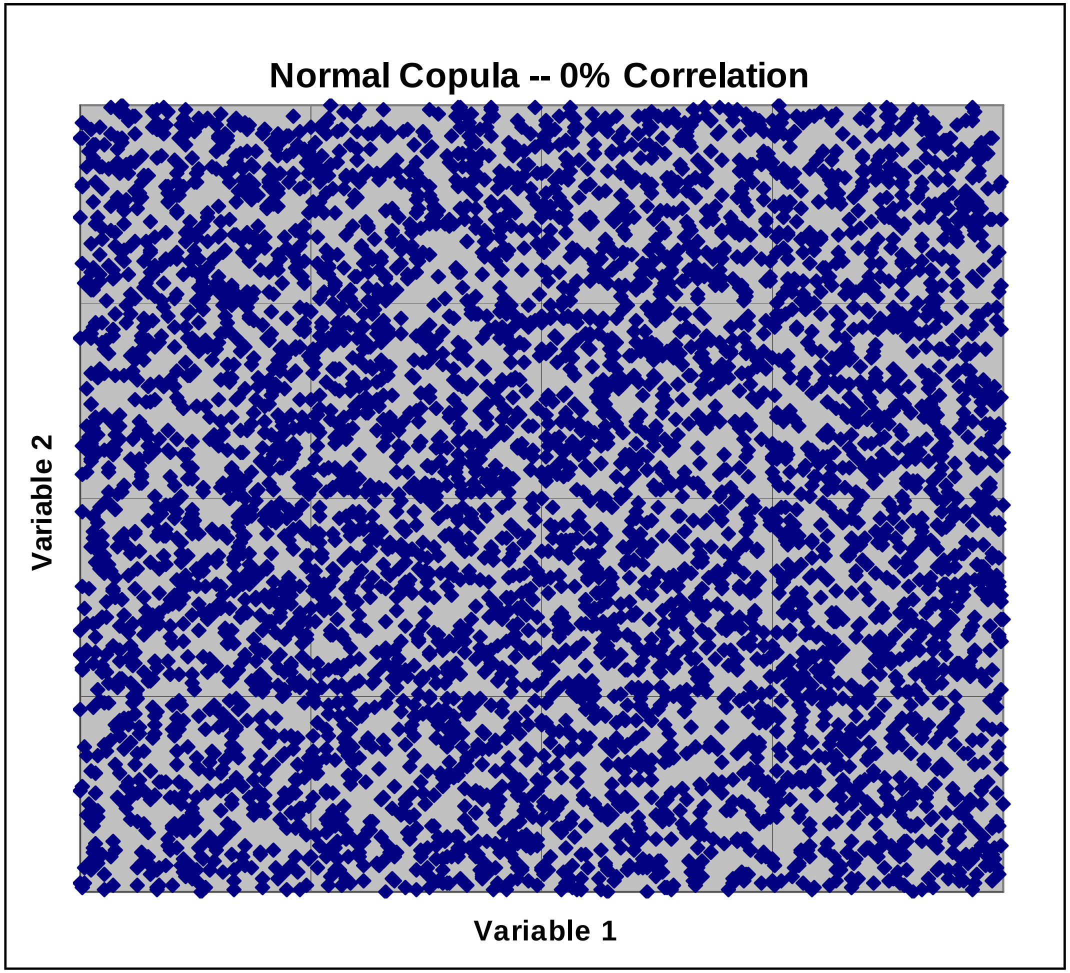
- Three popular copulas in actuarial use today are
  - The Normal copula
  - The Student-t copula
  - The Gumbel copula
- Copula required components (with the exception of Gumbel):
  - The marginal distributions of the individual lines
  - Correlations among these lines
- The Gumbel copula is different from the Normal and Student-t
  - It does not need a complete correlation matrix
  - Association is expressed by a single parameter applying to all lines
  - Upper tail dependence is strong while lower tail dependence always equals 0

The choice of the appropriate copula is a matter of judgment

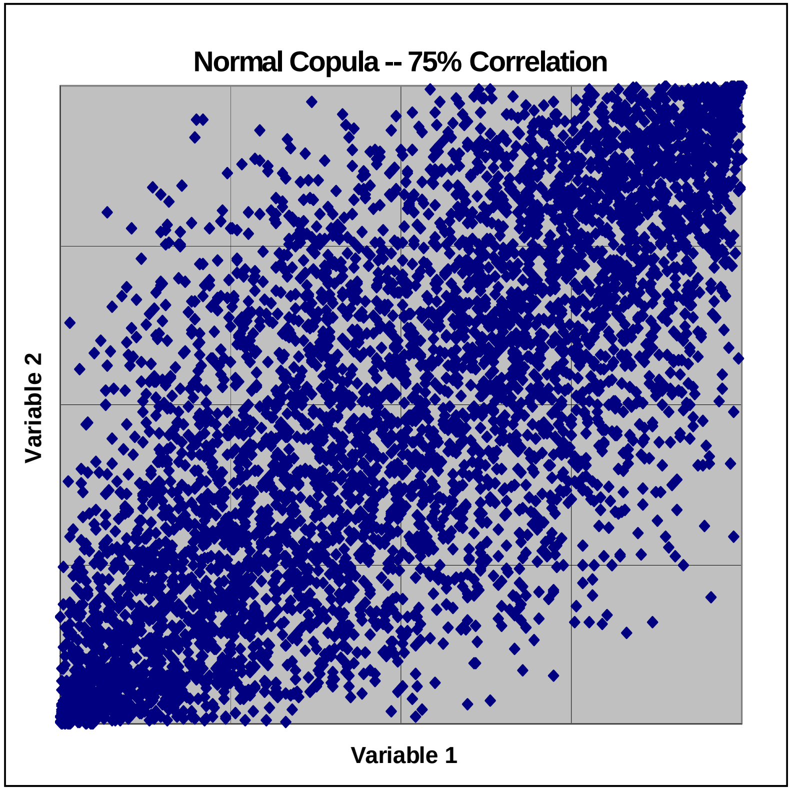


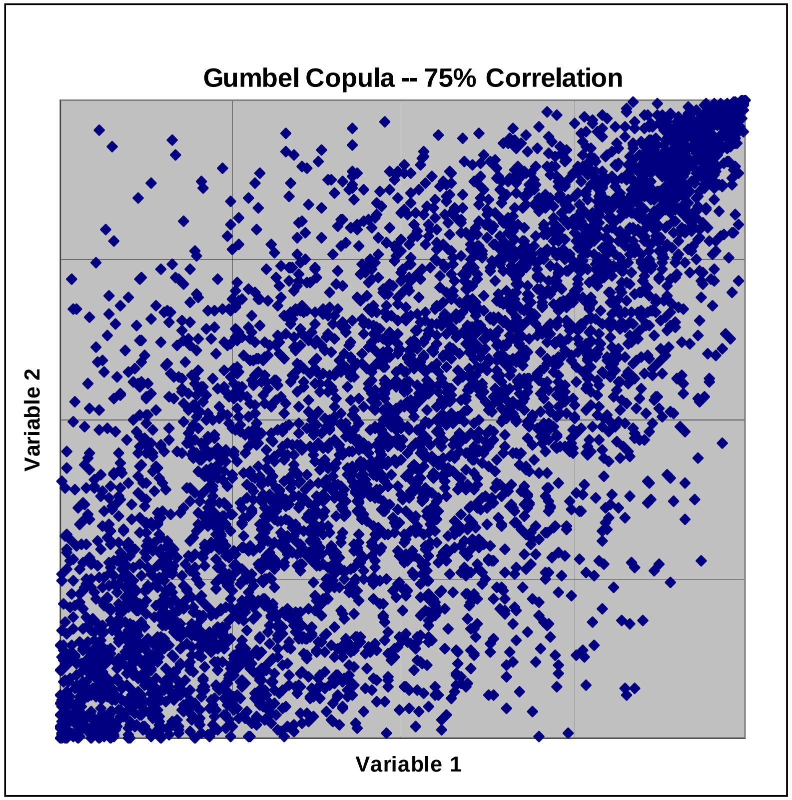
- The portfolio of liabilities can be stress-tested under varying copula assumptions

With independent variables results are not correlated

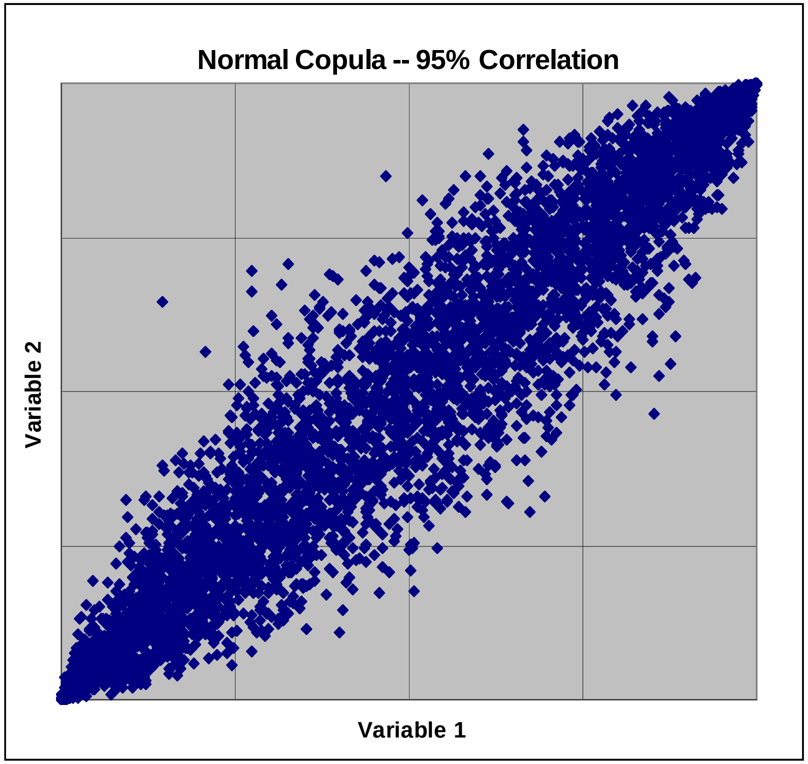


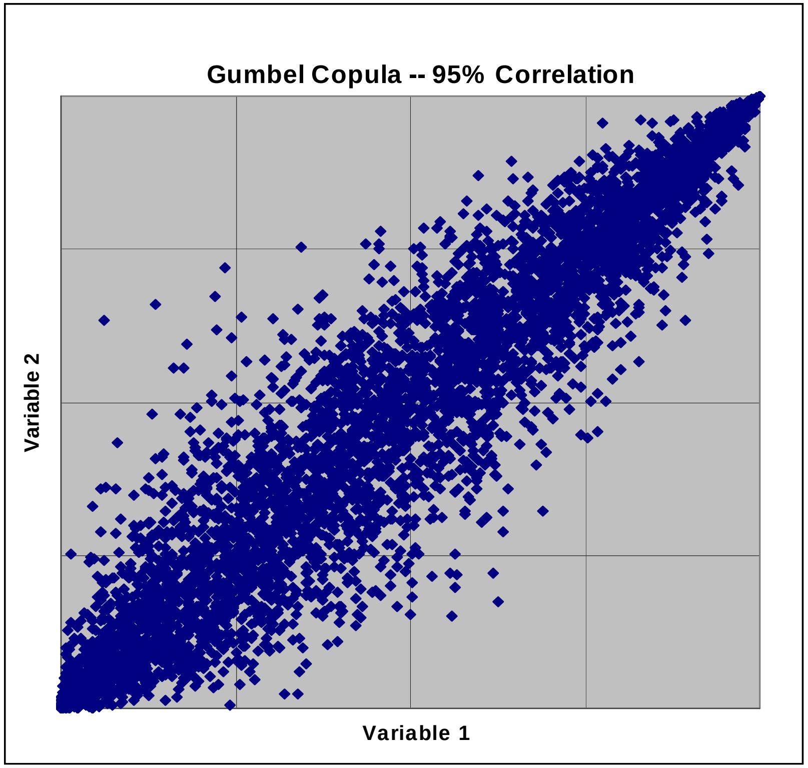
75% correlation: bad results in one line make it more likely to have bad results in the second line





The relationship is even more pronounced with 95% correlation





Questions?

?????????????????
???????????????
?????????????
???????????
?????????
???????
?????
???
?